

ABSTRACT OF THE DISCLOSURE

A method and system are provided for improving the transmission and storage of documents utilizing markup languages. The system converts the markup language tags to tokens and combines the tokens with the content information to create a token stream. The token stream is further compressed and stored on a device or transmitted over a communications network. The compressed token stream may then be decompressed and de-tokenized to return the markup language document to its original content form. The method may be performed on a computing device or over the transmission network.